Hidden Variables in Bipartite Networks

Maksim Kitsak[1] and Dmitri Krioukov[1]

[1] *Cooperative Association for Internet Data Analysis (CAIDA), University of California, San Diego (UCSD), 9500 Gilman Drive, La Jolla, CA 92093, USA*

(Dated: January 19, 2013)

We introduce and study random bipartite networks with hidden variables. Nodes in these networks are characterized by hidden variables which control the appearance of links between node pairs. We derive analytic expressions for the degree distribution, degree correlations, the distribution of the number of common neighbors, and the bipartite clustering coefficient in these networks. We also establish the relationship between degrees of nodes in original bipartite networks and in their unipartite projections. We further demonstrate how hidden variable formalism can be applied to analyze topological properties of networks in certain bipartite network models, and verify our analytical results in numerical simulations.

PACS numbers: 89.75.Hc, 05.45.Df, 64.60.Ak

I. INTRODUCTION

Bipartite networks are composed of two types of nodes with no links connecting nodes of the same type, see Fig. 1(a). Examples include recommendation systems [1], networks of collaborations [2] and metabolic reactions [3], gene regulatory networks [4], peer to peer networks [5], pollination networks [6], and many others [7]. Compared to traditional unipartite networks, less is known about the organizing principles determining the structure and evolution of bipartite networks, partly because only unipartite projections of bipartite networks are often considered. The unipartite projection accounts for connecting two nodes of one type by a link if these nodes share at least one neighbor of the other type, and then throwing out all nodes of this other type, see Figs. 1(b) and 1(c). Even though this procedure allows one to study bipartite networks using powerful tools developed for unipartite networks, the unipartite projections in most cases lead to significant loss of information, and to artificial inflation of the projected network with fully connected subgraphs [7, 8].

Nodes in real bipartite networks can often be characterized by a number of intrinsic attributes. For example, in recommendation networks, composed of consumer and product nodes, a consumer-product pair is connected if the consumer has purchased the product. Consumers can be characterized by their age, geographic location, income, sex, lifestyle, etc., while products have their type, price, quality, uniqueness, and other properties. Consumers do not buy products at random. Making their purchase decisions, consumers implicitly match their attributes with those of products. For example, a person with a higher income is more likely to purchase an expensive item, books in Italian are mostly purchased by people who speak Italian, consumers at a gas station tend to own a car, etc. Similar considerations apply to the formation of links between researchers and scientific projects, molecules and reactions in which they participate, and so forth.

The concept of hidden variables formalizes these observations as follows. Every node of each type in a bipartite network is assigned a number of hidden variables drawn from some distributions, and then every node pair of different types is connected with some probability which depends on the hidden variables of the two nodes. In this work we build the hidden variable formalism for bipartite networks, based on the formalism developed earlier for unipartite networks [9]. Specifically, in Section II we overview basic topological characteristics of bipartite networks. In Section III we define a general class of bipartite networks with hidden variables, and study analytically the topological properties of networks in this class. In Section IV we consider two specific examples of bipartite networks with hidden variables, uncorrelated and stratified bipartite networks, and confirm in simulations our analytical results for these networks. Section V summarizes the paper.

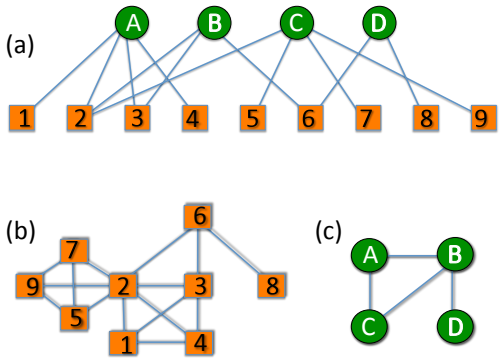

FIG. 1. (Color Online) A toy bipartite network and its unipartite projections. (a) Original bipartite network. We refer to the nodes of one type as top nodes (labeled by letters) and to the nodes of the other type as bottom nodes (labeled by numbers). Unipartite projections of the original network onto (b) bottom and (c) top domains. The top (bottom) nodes are connected in the projections if they have at least one common neighbor in the original network.

II. TOPOLOGICAL CHARACTERISTICS OF BIPARTITE NETWORKS

In this section we review some key relationships among the basic topological characteristics of bipartite networks.

Let the nodes of two different types be called top and bottom nodes, see Figs. 1(b) and 1(c). Similar to unipartite networks, the degree correlations in bipartite networks are defined by the number of links $E_{k\ell}$ between top and bottom nodes of degrees k and ℓ [10]. The correlation matrix $E_{k\ell}$ satisfies the following equations:

$$\sum_{\ell} E_{k\ell} = kN_k, \quad \sum_{k} E_{k\ell} = \ell M_\ell, \quad \sum_{k,\ell} E_{k\ell} = E, \quad (1)$$

where N_k and M_ℓ are the numbers of top and bottom nodes of degree k and ℓ, and E is the total number of links in the network. The joint degree distribution $P(k,\ell)$ is the normalized correlation matrix, i.e., the probability that a randomly chosen edge connects nodes of degrees k and ℓ:

$$P(k,\ell) = \frac{E_{k\ell}}{E}, \quad (2)$$

which contains all information needed to construct a network with a given degree distribution and correlations.

The top and bottom node degree distributions $P(k)$ and $P(\ell)$ can be obtained from Eq. (1):

$$P(k) = \frac{\overline{k}}{k}\sum_{\ell} P(k,\ell), \quad P(\ell) = \frac{\overline{\ell}}{\ell}\sum_{k} P(k,\ell). \quad (3)$$

The conditional probabilities $P(\ell|k)$ and $P(k|\ell)$ that an edge emanating from a k- or ℓ-degree node is connected to a node of degree ℓ or k are

$$P(\ell|k) = \frac{E_{k\ell}}{kN_k} = \frac{\overline{k}P(k,\ell)}{kP(k)}, \quad (4)$$

$$P(k|\ell) = \frac{E_{k\ell}}{\ell N_\ell} = \frac{\overline{\ell}P(k,\ell)}{\ell P(\ell)}. \quad (5)$$

To characterize degree correlations in unipartite networks, one often considers the average nearest neighbor degree (ANND), which is the average degree of the neighbors of all k-degree nodes [11]. The ANNDs for top and bottom nodes in a bipartite network are

$$\overline{\ell}_{nn}(k) = \sum_{\ell} \ell P(\ell|k), \quad \overline{k}_{nn}(\ell) = \sum_{k} kP(k|\ell). \quad (6)$$

In uncorrelated bipartite networks

$$P^{unc}(k,\ell) = \frac{kP(k)}{\overline{k}}\frac{\ell P(\ell)}{\overline{\ell}}. \quad (7)$$

As a result, $P(\ell|k)$ and $P(k|\ell)$ do not depend on k and ℓ, respectively:

$$P^{unc}(\ell|k) = \frac{\ell}{\overline{\ell}}P(\ell), \quad P^{unc}(k|\ell) = \frac{k}{\overline{k}}P(k), \quad (8)$$

and neither do the ANNDs:

$$\overline{\ell}^{unc}_{nn}(k) = \frac{\overline{\ell^2}}{\overline{\ell}}, \quad \overline{k}^{unc}_{nn}(\ell) = \frac{\overline{k^2}}{\overline{k}}. \quad (9)$$

Networks with increasing or decreasing ANNDs are called assortative or disassortative [12]. Some real bipartite networks have non-trivial degree correlation profiles, and therefore they can not be classified as either assortative or disassortative [7].

The standard clustering coefficient of node i quantifies how close i's neighbors are to forming a clique [13]:

$$c(i) = \frac{2}{k_i(k_i-1)}\sum_{j>k} e_{jk}, \quad (10)$$

where the summation is over all i's pairs of neighbors j and k, and e_{jk} is the adjacency matrix. Since in bipartite networks there are no loops of size 3, this clustering coefficient is zero for all nodes. Therefore, to assess the density of connections in a vicinity of a particular node, one has to analyze connections among its second nearest neighbors. There have been several attempts to generalize the clustering coefficient for bipartite networks using this idea [7, 14, 15]. Here we focus on the definition by Zhang et al [14]:

$$c_B(i) = \frac{\sum_{m>n} q_{imn}}{\sum_{m>n}(q_{imn} + k_m + k_n - 2\eta_{imn})}, \quad (11)$$

where $\sum_{m>n}$ goes over all pairs of i's neighbors, q_{imn} is the number of common neighbors between nodes m and n excluding i, k_m and k_n are the degrees of nodes m and n, and $\eta_{imn} = 1 + q_{imn} + e_{mn}$. The above definition may look cumbersome, but it has a simple interpretation. Let A_m and A_n be the sets of neighbors of nodes m and n excluding i. Then q_{imn} is the intersection of A_m and A_n, $q_{imn} = \|A_m \bigcap A_n\|$, while $q_{imn} + (k_m - \eta_{imn}) + (k_n - \eta_{imn}) = \|A_m \bigcup A_n\|$ is their union. Therefore, the bipartite clustering coefficient is simply

$$c_B(i) = \frac{\sum_{m>n}\|A_m \bigcap A_n\|}{\sum_{m>n}\|A_m \bigcup A_n\|}. \quad (12)$$

The ratio of the intersection and union of two sets is known as the Jaccard similarity coefficient [16]. The bipartite clustering coefficient, on the other hand, is given by the ratio of the sums of intersections and unions for all pairs of i's neighbors. Therefore, the bipartite clustering coefficient can be interpreted as a combined Jaccard similarity of i's neighbors. Regardless of the clustering definition details, nodes in real bipartite networks tend to be strongly clustered, as compared to nodes in their randomized counterparts with preserved degree distributions [7].

III. HIDDEN VARIABLE FORMALISM FOR BIPARTITE NETWORKS

We define the class of bipartite networks with hidden variables as follows:

(i) Each top and bottom nodes i and j are assigned hidden variables κ_i and λ_j drawn from probability distribution $\rho_t(\kappa)$ and $\rho_b(\lambda)$;

(ii) Each top-bottom node pair $\{i,j\}$ is connected with probability $r(\kappa_i,\lambda_j)$, $0 \le r(\kappa,\lambda) \le 1$.

The hidden variable formalism developed here is valid for both discrete and continuous variables. In the latter case, all sums must be replaced by integrals. We are primarily interested in the cases where the hidden variable distributions $\rho_t(\kappa)$ and $\rho_b(\lambda)$ are independent of the sizes of the top and bottom domains N and M. We also assume that in the thermodynamic limit of large N, M, these sizes are proportional to each other, $N \propto M$. For the sake of clarity we consider only one hidden variable per node. The generalization to several hidden variables per node is straightforward. We also drop indices in the top and bottom hidden variable distribution notations: $\rho_t(\kappa) \equiv \rho(\kappa)$ and $\rho_b(\lambda) \equiv \rho(\lambda)$.

A. Degree distributions

We first compute the most basic topological properties of the networks in the model—the degree distributions and average degrees. Due to the stochastic nature of connections between top and bottom nodes, we can not compute the degree of a top node with hidden variable κ deterministically. Instead, we can compute propagator $g(k|\kappa)$, which is the probability that a node with hidden variable κ ends up connecting to k bottom nodes. Similarly, propagator $f(\ell|\lambda)$ is the probability that a bottom node with hidden variable λ will be connected to ℓ top nodes. Propagators $g(k|\kappa)$ and $f(\ell|\lambda)$ are the main building blocks of the hidden variable formalism. As soon as we know $g(k|\kappa)$, for example, the average degree $\overline{k}(\kappa)$ of a top node with hidden variable κ, the degree distribution $P(k)$, and the average degree $\overline{k}$ in the top node domain are given by:

$$\overline{k}(\kappa) = \sum_k k g(k|\kappa), \tag{13}$$

$$P(k) = \sum_\kappa g(k|\kappa)\rho(\kappa), \tag{14}$$

$$\overline{k} = \sum_k k P(k) = \sum_\kappa \overline{k}(\kappa)\rho(\kappa), \tag{15}$$

while the corresponding expressions for bottom nodes can be obtained by an appropriate swap of notations.

To compute propagator $g(k|\kappa)$ we first compute partial propagator $g_i^{\lambda_i}(k_i|\kappa)$ defined as the probability that a top node with hidden variable κ ends up having k_i connections to bottom nodes with hidden variable λ_i. Since links between node pairs appear independently from one pair to another, $g_i^{\lambda_i}(k_i|\kappa)$ is given by the binomial distribution:

$$g_i^{\lambda_i}(k_i|\kappa) = C_{k_i}^{M_{\lambda_i}} \left[r(\kappa,\lambda_i)\right]^{k_i} \left[1 - r(\kappa,\lambda_i)\right]^{M_{\lambda_i} - k_i}, \tag{16}$$

where C_b^a is the binomial coefficient, and $M_{\lambda_i} \equiv M\rho(\lambda_i)$ is the total number of bottom nodes with hidden variable λ_i. The full propagator $g(k|\kappa)$ is then a convolution of partial propagators:

$$g(k|\kappa) = \sum_{\sum k_i = k} \prod_i g_i^{\lambda_i}(k_i|\kappa), \tag{17}$$

where the product is over the entire spectrum of hidden variables λ, while the summation is over the ensemble of all possible degrees k_i whose sum is k.

Since the full propagator is a convolution, its generating function $\hat{g}(z|\kappa)$ is a product of the generating functions $\hat{g}^\lambda(z|\kappa)$ for partial propagators:

$$\hat{g}(z|\kappa) = \prod_\lambda \hat{g}^\lambda(z|\kappa), \quad \text{where} \tag{18}$$

$$\hat{g}(z|\kappa) \equiv \sum_k g(k|\kappa) z^k, \tag{19}$$

$$\hat{g}^\lambda(z|\kappa) \equiv \sum_k g^\lambda(k|\kappa) z^k. \tag{20}$$

The generating function for binomial $g^\lambda(k|\kappa)$ is

$$\hat{g}^\lambda(z|\kappa) = (1 - z(1 - r(\kappa,\lambda)))^{M_\lambda}, \tag{21}$$

substituting which into Eq. (18) we obtain

$$\ln \hat{g}(z|\kappa) = M \sum_\lambda \rho(\lambda) \ln\left[1 - (1-z) r(\kappa,\lambda)\right]. \tag{22}$$

The average degree of nodes with hidden variable κ is given by the derivative of $\hat{g}(z|\kappa)$ at $z = 1$ [17], to confirm the obvious

$$\overline{k}(\kappa) = M \sum_\lambda \rho(\lambda) r(\kappa,\lambda), \tag{23}$$

while higher moments of $g(k|\kappa)$ can be computed by taking higher order derivatives of the generating function. Eq. (23) yields the average degree in the entire top node domain

$$\overline{k} = \sum_k \overline{k}(\kappa)\rho(\kappa) = M \sum_{\kappa,\lambda} \rho(\kappa)\rho(\lambda) r(\kappa,\lambda), \tag{24}$$

and the expected total number of links in the network

$$E = N\overline{k} = M\overline{\ell} = NM \sum_{\kappa,\lambda} \rho(\kappa)\rho(\lambda) r(\kappa,\lambda). \tag{25}$$

It is evident from the last equation that to end up with a sparse bipartite network, $E \propto N \propto M$, the connection probability $r(\kappa,\lambda)$ must be of the form

$$r(\kappa,\lambda) \propto \hat{r}(\kappa,\lambda)/M, \tag{26}$$

where $\hat{r}(\kappa,\lambda)$ is independent of M. Therefore, for large sparse networks we can expand the logarithm in Eq. (22) in powers of $r(\kappa,\lambda)$ to finally obtain, in the first order,

$$\ln \hat{g}(z|\kappa) \approx (z-1) \sum_\lambda \rho(\lambda)\hat{r}(\kappa,\lambda), \tag{27}$$

$$g(k|\kappa) = e^{-\overline{k}(\kappa)} \left[\overline{k}(\kappa)\right]^k / k!, \tag{28}$$

which we can use to compute the degree distribution in Eq. (14). Propagator $f(\ell|\lambda)$ and degree distribution $P(\ell)$ for bottom nodes can be obtained from Eqs. (28) and (14) by swapping $\kappa \to \lambda$ and $k \to \ell$.

The Poisson form of the propagator $g(k|\kappa)$, given by Eq. (28), implies that

$$\overline{k^2}(\kappa) = \left[\overline{k}(\kappa)\right]^2 + \overline{k}(\kappa). \tag{29}$$

Furthermore, Eq. (14) allows us to obtain the second moment of the degree distribution:

$$\overline{k^2} = \sum_k k^2 P(k) = \sum_\kappa \left[\overline{k}(\kappa)\right]^2 \rho(\kappa) + \sum_\kappa \overline{k}(\kappa)\rho(\kappa) \tag{30}$$

B. Unipartite projection

Next we establish the connection between the degrees of nodes in a bipartite network and in its unipartite projections, often considered in the literature. In the top unipartite projection, two top nodes are connected if they have at least one common bottom neighbor in the bipartite network. Therefore, we first compute the probability $p_0(\kappa_1, \kappa_2)$ that two top nodes with hidden variables κ_1 and κ_2 do not have any common bottom neighbors in the bipartite network. This probability is

$$p_0(\kappa_1, \kappa_2) = \prod_i \left[1 - r(\kappa_1, \lambda_i) r(\kappa_2, \lambda_i)\right], \tag{31}$$

where the product is over all the bottom nodes. Taking the logarithm on both sides, we get

$$\ln p_0(\kappa_1, \kappa_2) = M \sum_\lambda \rho(\lambda) \ln\left[1 - r(\kappa_1, \lambda) r(\kappa_2, \lambda)\right], \tag{32}$$

and the probability $p_u(\kappa_1, \kappa_2) = 1 - p_0(\kappa_1, \kappa_2)$ that two top nodes with hidden variables κ_1 and κ_2 are connected in the unipartite projection is simply

$$p_u(\kappa_1, \kappa_2) = 1 - \exp\{M \sum_\lambda \rho(\lambda) \ln\left[1 - r(\kappa_1, \lambda) r(\kappa_2, \lambda)\right]\}. \tag{33}$$

In sparse networks we use Eq. (26) to approximate $p_u(\kappa_1, \kappa_2)$ as

$$p_u(\kappa_1, \kappa_2) \approx M \sum_\lambda \rho(\lambda) r(\kappa_1, \lambda) r(\kappa_2, \lambda). \tag{34}$$

Next we find propagator $p(k_u|\kappa)$, the conditional probability that a top node with hidden variable κ has k_u connections in the unipartite projection. The derivation is similar to the derivation of propagator $g(k|\kappa)$ for the bipartite network. We first define partial propagator $p_i^{\kappa_i'}(n_i|\kappa)$, the probability that a top node with hidden variable κ is connected in the unipartite projection to n_i nodes with hidden variable κ_i'. Equation (34) indicates that a node with hidden variable κ is equally likely to be connected in the unipartite projection to any of $N_{\kappa_i'}$ nodes with hidden variable κ_i', where $N_{\kappa_i'} = N\rho(\kappa_i')$ is the number of top nodes with hidden variable κ_i'. If $n_i \ll M$, we can assume that the links in the unipartite projection are independent, leading to binomial $p_i^{\kappa_i'}(n_i|\kappa)$:

$$p_i^{\kappa_i'}(n_i|\kappa) = C_{n_i}^{N_{\kappa_i'}} \left[p_u(\kappa, \kappa_i')\right]^{n_i} \left[(1 - p_u(\kappa, \kappa_i'))\right]^{N_{\kappa_i'} - n_i}. \tag{35}$$

Similar to Eq. (17), $p(k_u|\kappa)$ is then a convolution

$$p(k_u|\kappa) = \sum_{\sum n_i = k_u} \prod_i p_i^{\kappa_i'}(n_i|\kappa), \tag{36}$$

and its generating function $\hat{p}(z|\kappa) = \sum_{k_u} p(k_u|\kappa) z^{k_u}$ is

$$\ln \hat{p}(z|\kappa) = N \sum_{\kappa'} \rho(\kappa') \ln\left[1 - (1 - z) p_u(\kappa, \kappa')\right]. \tag{37}$$

Therefore if $p_u(\kappa, \kappa')$ scales as

$$p_u(\kappa, \kappa') \sim \frac{1}{N^a}, \tag{38}$$

with $a \geq 1$, then similar to the bipartite case, propagator $p(k_u|\kappa)$ is approximately the Poisson distribution:

$$p(k_u|\kappa) \approx e^{-\overline{k_u}(\kappa)} \left[\overline{k_u}(\kappa)\right]^{k_u} / k_u!. \tag{39}$$

The average degree $\overline{k_u}(\kappa)$ of nodes with hidden variable κ in the unipartite projection is given by the first derivative of the generating function $\hat{p}(z|\kappa)$ at $z = 1$ to yield the obvious

$$\overline{k_u}(\kappa) = N \sum_{\kappa'} \rho(\kappa') p_u(\kappa, \kappa'), \tag{40}$$

which for sparse networks using Eq. (34) transforms to:

$$\overline{k_u}(\kappa) = NM \sum_{\lambda, \kappa'} \rho(\lambda)\rho(\kappa') r(\kappa, \lambda) r(\kappa', \lambda) \tag{41}$$

$$= M \sum_\lambda \overline{\ell}(\lambda)\rho(\lambda) r(\kappa, \lambda), \tag{42}$$

where $\overline{\ell}(\lambda)$ is the average degree of bottom nodes with hidden variable λ in the bipartite network. The average degree in the entire top unipartite projection is then

$$\overline{k_u} = \sum_\kappa \rho(\kappa) \overline{k_u}(\kappa) = \frac{M}{N} \sum_\lambda \rho(\lambda) \left[\overline{\ell}(\lambda)\right]^2. \tag{43}$$

Finally, the degree distribution in the unipartite projection is

$$P(k_u) = \sum_\kappa p(k_u|\kappa)\rho(\kappa). \tag{44}$$

C. Number of common neighbors

The common neighbor statistics is useful in many applications, such as node similarity estimation [18] and link prediction [19]. We compute the probability that two top nodes with hidden variables κ_1 and κ_2 have m common bottom neighbors. This probability can be calculated as

$$P_{\kappa_1,\kappa_2}(m) = \sum_{\sum m_i = m} \prod_i p_{\kappa_1,\kappa_2}(m_i|\lambda_i), \tag{45}$$

where $p_{\kappa_1,\kappa_2}(m_i|\lambda_i)$ is the probability that two top nodes with κ_1 and κ_2 have m_i common bottom neighbors with λ_i, and the product is over the entire range of λ_i, while the summation is over all possible combinations of m_i adding up to m.

Consider two nodes with hidden variables κ_1 and κ_2. Each common neighbor of the two nodes with κ_1 and κ_2 appears independently with probability

$$\tilde{r}_\lambda(\kappa_1,\kappa_2) = r(\kappa_1,\lambda) r(\kappa_2,\lambda), \tag{46}$$

where λ is the hidden variable of the common neighbor. Therefore, $p_{\kappa_1,\kappa_2}(m|\lambda)$ is also binomial:

$$p_{\kappa_1,\kappa_2}(m|\lambda) = C_m^{M_\lambda} \left[\tilde{r}_\lambda(\kappa_1,\kappa_2))\right]^m \left[1-\tilde{r}_\lambda(\kappa_1,\kappa_2)\right]^{M_\lambda - m}, \tag{47}$$

and the corresponding generating function is given by

$$\widehat{p}_{\kappa_1,\kappa_2}(z|\lambda) = \left[1-(1-z)\tilde{r}_\lambda(\kappa_1,\kappa_2)\right]^{M_\lambda}. \tag{48}$$

Since $P_{\kappa_1,\kappa_2}(m)$ is given by a convolution, its generation function is

$$\widehat{P}_{\kappa_1,\kappa_2}(z) = \prod_i \widehat{p}_{\kappa_1,\kappa_2}(z|\lambda_i). \tag{49}$$

Combining the last two equations, we get

$$\ln \widehat{P}_{\kappa_1,\kappa_2}(z) = M \sum_\lambda \rho(\lambda) \ln\left[1-(1-z)\tilde{r}_\lambda(\kappa_1,\kappa_2)\right]. \tag{50}$$

To compute the average number of common neighbors between top nodes with κ_1 and κ_2 we evaluate the derivative of $\widehat{P}_{\kappa_1,\kappa_2}(z)$ with respect to z at $z=1$:

$$\overline{m}(\kappa_1,\kappa_2) = M \sum_\lambda \rho(\lambda) \tilde{r}_\lambda(\kappa_1,\kappa_2). \tag{51}$$

The generating function for the common neighbor distribution has the same structure as $\hat{g}(z|k)$. Therefore, the closed form of $P_{\kappa_1,\kappa_2}(m)$ in the sparse network approximation is given by

$$P_{\kappa_1,\kappa_2}(m) \approx e^{-\overline{m}(\kappa_1,\kappa_2)} \left[\overline{m}(\kappa_1,\kappa_2)\right]^m / m!. \tag{52}$$

D. Degree correlations

The degree correlations in bipartite networks are fully described by conditional probabilities $P(\ell|k)$ and $P(k|\ell)$ in Eqs. (4,5). In order to calculate $P(\ell|k)$ we need to define the related conditional probability $\rho(\lambda|\kappa)$ that an edge outgoing from a top node with hidden variable κ is connected to a bottom node with hidden variable λ. Then, $P(\ell|k)$ can be written as

$$P(\ell|k) = \sum_{\kappa,\lambda} f(\ell-1|\lambda)\rho(\lambda|\kappa) g^*(\kappa|k), \tag{53}$$

where $f(\ell-1|\lambda)$ is the conditional probability that a bottom node with hidden variable λ ends up having degree ℓ (one connection is already taken into account by the conditional edge), while the inverse propagator $g^*(k|\kappa)$ is the probability that a top node of degree k has hidden variable κ. This inverse propagator is given by the Bayes' formula [20]

$$P(k) g^*(\kappa|k) = \rho(\kappa) g(k|\kappa), \tag{54}$$

using which we write

$$P(\ell|k) = \frac{1}{P(k)} \sum_{\kappa,\lambda} \rho(\kappa)\rho(\lambda|\kappa) f(\ell-1|\lambda) g(k|\kappa). \tag{55}$$

To determine $\rho(\lambda|\kappa)$ we note that the conditional probability that an edge is connected to a bottom node with λ, given that this edge is connected to a top node with κ, is proportional to the density of bottom nodes $\rho(\lambda)$ and the connection probability $r(\kappa,\lambda)$,

$$\rho(\lambda|\kappa) \propto \rho(\lambda) r(\kappa,\lambda). \tag{56}$$

Taking into account the normalization condition $\sum_\lambda \rho(\lambda|\kappa) = 1$, we get

$$\rho(\lambda|\kappa) = \frac{\rho(\lambda) r(\kappa,\lambda)}{\sum_{\lambda'} \rho(\lambda') r(\kappa,\lambda')}. \tag{57}$$

Using Eqs. (55-57) we obtain the final expression for the top ANND statistics:

$$\overline{\ell}_{nn}(k) = \sum_\ell \ell P(\ell|k) = 1 + \frac{1}{P(k)} \sum_\kappa \overline{\ell}_{nn}(\kappa)\rho(\kappa) g(k|\kappa), \tag{58}$$

where $\overline{\ell}_{nn}(\kappa)$ is the average nearest neighbor degree of top nodes with hidden variable κ:

$$\overline{\ell}_{nn}(\kappa) = \sum_\lambda \overline{\ell}(\lambda)\rho(\lambda|\kappa). \tag{59}$$

E. Bipartite clustering coefficient

Finally we derive the bipartite clustering coefficient as defined by P. Zhang et al [14]. Other variations of the bipartite clustering coefficient can be computed in a similar manner.

The bipartite clustering coefficient of top node i, given by Eq. (11), can be written as

$$c_B(i) = \frac{\sum_{j>l}(m_{jl}-1)}{\sum_{j>l}(k_j+k_l-m_{jl}-1)}, \tag{60}$$

where m_{jl} is the number of common neighbors between bottom nodes j and l, while k_j and k_l are their degrees. Since the summations in the numerator and denominator are performed independently, we can estimate the average bipartite clustering coefficient of top nodes with hidden variable κ by calculating the ensemble averages of the numerator and denominator. The details are in the Appendix, while the answer is

$$\overline{c_B}(\kappa) = \frac{\sum_{\lambda_1,\lambda_2}\rho(\lambda_1|\kappa)\rho(\lambda_2|\kappa)\overline{m}(\lambda_1,\lambda_2)}{2\bar{\ell}_{nn}(\kappa)-\sum_{\lambda_1,\lambda_2}\rho(\lambda_1|\kappa)\rho(\lambda_2|\kappa)\overline{m}(\lambda_1,\lambda_2)}, \tag{61}$$

where $\rho(\lambda|\kappa)$ is the conditional probability that an edge connected to a top node with hidden variable κ is also connected to a bottom node with hidden variable λ, $\overline{m}(\lambda_1,\lambda_2)$ is the average number of common neighbors between two bottom nodes with hidden variables λ_1 and λ_2, and $\bar{\ell}_{nn}(\kappa)$ is the average nearest neighbor degree of top nodes with hidden variable κ. The average bipartite clustering coefficient of top nodes with degrees $k \geq 2$ can be expressed in terms of $\overline{c_B}(\kappa)$ as

$$\overline{c_B}(k) = \frac{1}{P(k)}\sum_{\kappa}\rho(\kappa)g(k|\kappa)\overline{c_B}(\kappa), \tag{62}$$

while the average bipartite clustering coefficient in the top node domain is simply

$$\overline{c_B} = \sum_{\kappa}\rho(\kappa)\overline{c_B}(\kappa) = \sum_{k}P(k)\overline{c_B}(k). \tag{63}$$

IV. EXAMPLES OF BIPARTITE NETWORKS WITH HIDDEN VARIABLES

Having the general formalism in place, we next consider a couple of examples of bipartite networks with hidden variables. The first example of uncorrelated networks is fairly standard. The second one, stratified networks, is more unusual.

A. Uncorrelated Bipartite Networks

Consider a random bipartite network composed of nodes with degrees $\{k_i\}$ and $\{\ell_j\}$ drawn from distributions $P(k)$ and $P(\ell)$. If nodes in the network are connected at random, then two randomly chosen nodes with degrees k and ℓ are connected with probability $p = k\ell/E$, where E is the total number of links in the network.

Similar random uncorrelated networks can be constructed in the hidden variable formalism. Consider a network with hidden variables drawn from distributions $\rho(\kappa)$ and $\rho(\lambda)$, in which node pairs are connected with probability proportional to the product of nodes' hidden variables:

$$r(\kappa,\lambda) = \frac{\kappa\lambda}{C}, \tag{64}$$

where C is some normalization constant. The above form of $r(\kappa,\lambda)$ implies that the hidden variable of a node can be regarded as its target or expected degree. Indeed, if we choose $C = \overline{\lambda}M$, then a top node with hidden variable κ gets κ connections on average

$$\overline{k}(\kappa) = M\sum_{\lambda}\rho(\lambda)r(\kappa,\lambda) = \kappa. \tag{65}$$

Since the assumption of a sparse network, given by Eq. (26) holds here, propagator $g(k|\kappa)$ is given by the Poisson distribution:

$$g(k|\kappa) = e^{-\kappa}\kappa^k/k!, \tag{66}$$

and using Eqs. (14) and (29) one can obtain

$$\overline{k^2} = \overline{\kappa^2} + \overline{\kappa}. \tag{67}$$

One type of nodes in real bipartite networks is often characterized by scale-free degree distributions, while degree of nodes of the other type can follow either fat-tailed or poissonian distributions [8]. Our uncorrelated formalism can account for both options. The former case is actually simpler, and the properties of top and bottom nodes can be obtained from each other via a simple swap of notations. Therefore below we consider the latter case, which is more typical for real networks.

Specifically, let κ be power-law distributed:

$$\rho(\kappa) = (\gamma-1)k_0^{\gamma-1}\kappa^{-\gamma}, \tag{68}$$

where power-law exponent γ and minimum expected degree κ_0 are parameters of the distribution. The resulting degree distribution of the top node domain is given by Eqs. (14) and (28), which yield

$$P(k) = (\gamma-1)\kappa_0^{\gamma-1}\frac{\Gamma[k-\gamma+1,\kappa_0]}{\Gamma[k+1]}, \tag{69}$$

where $\Gamma[x,s]$ is the incomplete gamma function. In the large k limit we can approximate $P(k)$ by

$$P(k) \approx (\gamma-1)\kappa_0^{\gamma-1}k^{-\gamma}. \tag{70}$$

We note that the distribution $P(k)$ of top node degrees does not depend on a specific form of the hidden variable distribution $\rho(\lambda)$ in the bottom node domain. Let the latter be a delta function $\rho(\lambda) = \delta(\lambda-\lambda_0)$, meaning that all bottom nodes have the same value of their hidden variable equal to λ_0. Then using the same Eqs. (14,28) swapped for the bottom nodes, we immediately conclude

that the distribution of bottom node degrees is poissonian:

$$P(\ell) = e^{-\lambda_0}\lambda_0^{\ell}/\ell!. \tag{71}$$

We now turn our attention to the unipartite projections. We first consider the top node projection. We use Eq. (42) to compute the average degree of κ-nodes:

$$\overline{k}_u(\kappa) = \kappa\lambda_0. \tag{72}$$

Therefore the average degree in the top unipartite projection is

$$\overline{k}_u = \overline{\kappa}\lambda_0. \tag{73}$$

The degree distribution in the projection are given by Eqs. (44) and (39):

$$P(k_u) = (\gamma-1)[\kappa_0\lambda_0]^{\gamma-1}\frac{\Gamma[k_u-\gamma+1,\kappa_0\lambda_0]}{\Gamma[k_u+1]}, \tag{74}$$

that is, this distribution is also a power law,

$$P(k_u) \sim (\gamma-1)[\kappa_0\lambda_0]^{\gamma-1}k_u^{-\gamma}, \tag{75}$$

and the exponent of this power law is equal to the exponent of the top power-law degree distribution in the original bipartite network.

In the bottom unipartite projection, the average node degrees are obtained in a similar manner to yield

$$\overline{\ell_u} = \overline{\ell_u}(\lambda) = \lambda_0\frac{\overline{\kappa^2}}{\overline{\kappa}}. \tag{76}$$

For $\gamma \leq 3$, $\overline{\kappa^2}$ depends on N, and diverges in the thermodynamic limit. Therefore connection probability $p(\lambda_1,\lambda_2)$ does not satisfy the condition of Eq. (38), and we can not approximate the degree distribution in the bottom domain by Eq. (44) with poissonian $p(k_u|\kappa)$ in Eq. (39). However, if $\gamma > 3$, then $\overline{\kappa^2}$ is finite in the thermodynamic limit, and the degree distribution is given by

$$P(\ell_u) = e^{-\overline{\ell_u}}[\overline{\ell_u}]^{\ell}/\ell!. \tag{77}$$

As far as correlations are concerned, the conditional hidden variable distributions are

$$\rho(\lambda|\kappa) = \delta(\lambda-\lambda_0), \tag{78}$$

$$\rho(\kappa|\lambda) = \frac{\kappa}{\overline{\kappa}}\rho(\kappa), \tag{79}$$

leading to the following expression for the top and bottom ANNDs given by Eq. (58):

$$\overline{\ell}_{nn}(k) = 1+\lambda_0, \tag{80}$$

$$\overline{k}_{nn}(\ell) = 1+\frac{\overline{\kappa^2}}{\overline{\kappa}} = \frac{\overline{k^2}}{\overline{k}}. \tag{81}$$

The average number of common neighbors is given by Eq. (51) yielding, for top and bottom nodes,

$$\overline{m}(\kappa_1,\kappa_2) = \frac{\kappa_1\kappa_2}{M}, \tag{82}$$

$$\overline{m}(\lambda_0,\lambda_0) = \frac{\lambda_0^2}{N}\frac{\overline{\kappa^2}}{\overline{\kappa}^2}. \tag{83}$$

Finally, to compute clustering, we insert the expressions for the average number of common neighbors (82,83), ANNDs (80,81), and conditional distributions (78,79) into Eq. (61), and obtain the average bipartite clustering coefficient for top and bottom nodes:

$$\overline{c}_B(\kappa) = \frac{(\lambda_0^2\overline{\kappa^2})/(N\overline{\kappa}^2)}{2\lambda_0-(\lambda_0^2\overline{\kappa^2})/(N\overline{\kappa}^2)} \approx \frac{\lambda_0}{2N}\frac{\overline{\kappa^2}}{\overline{\kappa}^2}, \tag{84}$$

$$\overline{c}_B(\lambda) = \frac{(\overline{\kappa^2})^2/(M\overline{\kappa}^2)}{2\overline{\kappa^2}/\overline{\kappa}-(\overline{\kappa^2})^2/(M\overline{\kappa}^2)} \approx \frac{\overline{\kappa^2}}{2M\overline{\kappa}}. \tag{85}$$

We observe that the clustering coefficient of a node does not depend on its hidden variable in either case, i.e., that it is constant. This constant decreases as the network sizes N, M increase, and vanishes in the thermodynamic limit.

To test our analytical results we perform simulations, setting $N = 2M$, $N = 2\times10^5$, $\gamma = 2.5$, $\kappa_0 = 1$, and $\lambda_0 = 6$ to satisfy $N\overline{k} = M\overline{\ell}$. The degree distributions in the top and bottom domains as well as in their unipartite projections are shown in Fig. (2). The degree distribution of top nodes in the original bipartite network, and in its top unipartite projection both follow a power law with the same exponent $\gamma = 2.5$, see Fig. 2(a). As seen in Fig. 2(b), the degree distribution in the bottom node domain is well approximated by a Poisson distribution. On the other hand, due to the divergent behavior of the second moment of the top degree distribution $\overline{\kappa^2}$, the degree distribution in the bottom unipartite projection seems to follow a truncated power-law. The measured values of $\overline{k_u} = 20.0$ and $\overline{\ell_u} \approx 119$ are in good agreement with Eqs. (73,76) since $\overline{\kappa} = 2.85$ and $\overline{\kappa^2} = 62$ for the selected parameters.

In Fig. 3(a) we plot the ANNDs, and confirm that they are independent of node degrees as Eqs. (80,81) predict for uncorrelated networks.

To test the dependence of the average bipartite clustering coefficient on the network size, we generate a number of uncorrelated bipartite network of different sizes and values of γ. While sampling hidden variables κ for top nodes, we impose the cutoff of $\kappa_{max} \sim N^{1/2}$ to avoid structural degree correlations [21]. Therefore, $\overline{\kappa^2} \sim N^{(3-\gamma)/2}$, and the average bipartite clustering coefficient scales as $\overline{c}_B \sim N^{-\delta}$ with $\delta = (\gamma-1)/2$ for $2<\gamma<3$. In Fig. 3(b) we confirm this scaling. The figure shows the measured bipartite clustering coefficients as a function of N for different values of γ.

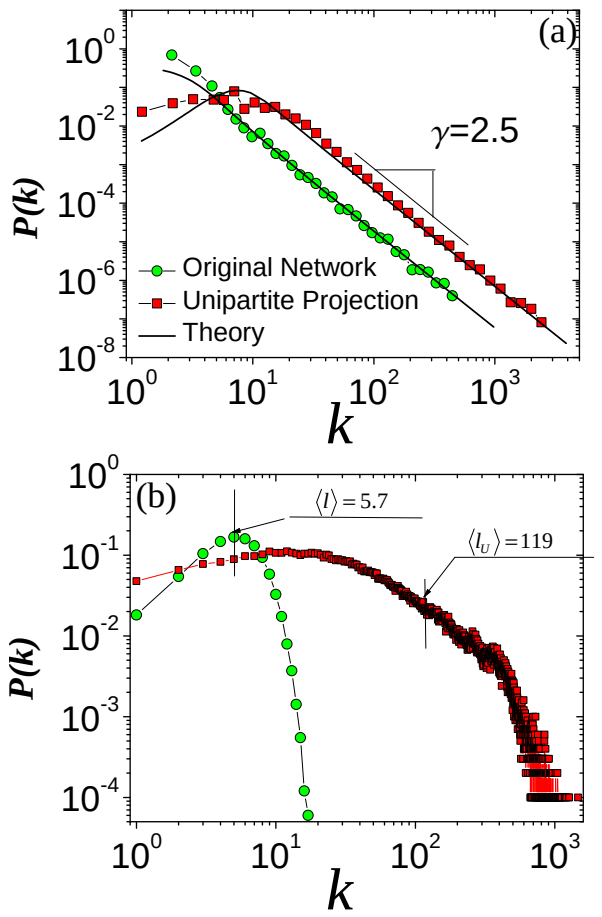

FIG. 2. (Color Online) Degree distributions in a random uncorrelated bipartite network. (a) Degree distributions in the top domain (green circles) and top unipartite projection (red squares). The solid lines are the analytical predictions from Eqs. (69,74). (b) Degree distributions in the bottom domain (green circles) and bottom unipartite projection (red squares). Both plots correspond to the model with $N = 2 \times 10^5$, $M = 10^5$, $\gamma = 2.5$, $\kappa_0 = 1$, and $\lambda_0 = 6$.

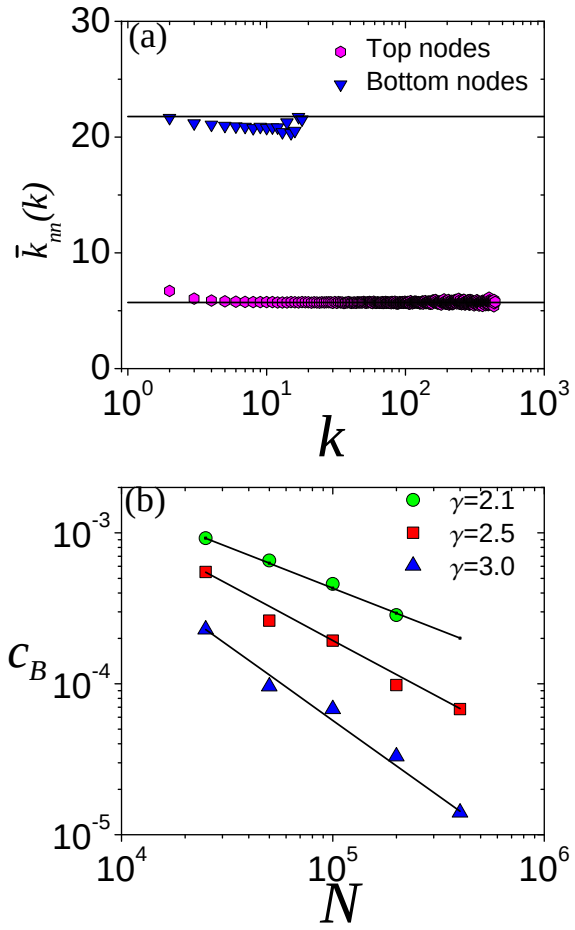

FIG. 3. (Color Online) (a) The average nearest neighbor degrees of top (blue triangles) and bottom (magenta hexagons) nodes in an uncorrelated bipartite network with $N = 2 \times 10^5$, $M = 10^5$, $\gamma = 2.5$, $\kappa_0 = 1$, and $\lambda_0 = 6.0$. The solid lines are the analytical predictions in Eqs. (80,81). (b) The average bipartite clustering coefficient for top nodes in uncorrelated bipartite networks as a function of network size N for $\gamma = 2.1$, $\gamma = 2.5$ and $\gamma = 3.0$. The solid lines are the theoretical predictions of $\bar{c}_B \sim N^{-\delta}$ with $\delta = (\gamma - 1)/2$.

B. Stratified Bipartite Networks

The original stratified unipartite network model was considered by Leicht et al [18]. In this model, N nodes are assigned random integer *ages* $t_i = 1, \ldots, t_{max}$ with uniform probability, and then links are created between node pairs with probability

$$P(\Delta t) = p_0 e^{-a|\Delta t|}, \tag{86}$$

where p_0 and a are model parameters. The motivation for this model in [18] was to have a simplified social model in which individuals preferably connect to other individuals of similar age. The stratified model was used in [18] to test the ability of different node similarity measures to infer relative node ages.

Here we generalize the stratified network model as follows. The networks in the model consist of N top and M bottom nodes. All nodes are assigned hidden variables κ and λ drawn from the continuous uniform distribution on interval $[0, T]$, $\rho(\kappa) = \rho(\lambda) = 1/T$. To eliminate finite size effects we impose the periodic boundary condition, meaning that nodes are uniformly scattered along a circle, and their hidden variables are their angular coordinates if we set $T = 2\pi$. To simplify the calculations we use the squared distances in the connection probability function:

$$r(\kappa, \lambda) = r_0 e^{-a\|\lambda - \kappa\|^2}, \tag{87}$$

where $\|\lambda - \kappa\|$ is the angular distance between λ and κ:

$$\|\lambda - \kappa\| = \pi - |\pi - |\lambda - \kappa||. \tag{88}$$

We first calculate the degree distributions for the top nodes. Due to the uniform distribution of hidden variables, the expected degree of a node is independent of its

hidden variable κ. Using Eqs. (23) and (24) we obtain

$$\overline{k} = \overline{k}(\kappa) = \frac{Mr_0}{2\sqrt{\pi a}}\mathrm{Erf}(\pi\sqrt{a}), \tag{89}$$

where $\mathrm{Erf}(x)$ is the error function. For $\overline{k}$ to be independent of network size, we must set $r_0/\sqrt{a} \sim 1/M$. Another natural choice would be to constraint $r_0 = M^{-1}$, but this choice would lead to bipartite clustering coefficients dependent on the network size. Constant bipartite clustering can be instrumented by setting

$$r_0 = 1, \quad \text{and} \quad a = \widetilde{a}M^2, \tag{90}$$

where $\widetilde{a}$ is a parameter controlling the average degree in the network. With the above choice of parameters Eq. (89) simplifies to

$$\overline{k} = \overline{k}(\kappa) \approx \frac{1}{2\sqrt{\pi\widetilde{a}}}. \tag{91}$$

Similarly, the average degree in the bottom node domain is given by

$$\overline{\ell} = \overline{\ell}(\lambda) = \frac{N}{M}\overline{k}. \tag{92}$$

Since connection probability $r(\kappa, \lambda)$ does not scale as M^{-1}, propagator $g(k|\kappa)$ is not given by Eq. (28). Instead we have to use Eq. (22) to compute the propagator, yielding

$$\hat{g}(z|\kappa) = e^{-\overline{k}\mathrm{Li}_{3/2}(1-z)}, \tag{93}$$

where $\mathrm{Li}_n(x)$ is the polylogarithm. Equation (93) can be used to calculate higher moments of the degree distribution. For example, the second moment is

$$\overline{k^2} = \overline{k}^2 + \overline{k}(1 - \frac{1}{\sqrt{2}}). \tag{94}$$

That is, similar to the Poisson distribution, the standard deviation of $g(k|\kappa)$ is

$$\sigma = \sqrt{\overline{k^2} - \overline{k}^2} \propto \sqrt{\overline{k}}. \tag{95}$$

According to Eq. (59), the average nearest neighbor degree is independent of the node's hidden variable:

$$\overline{\ell}_{nn}(\kappa) = \overline{\ell}, \tag{96}$$

because node degrees are not correlated with their hidden variables, see Eq. (91). Therefore, despite strong correlation between hidden variables of connected nodes, there are no degree correlations. The ANND can be obtained by inserting $\overline{\ell}_{nn}(\kappa)$ from Eq. (96) into Eq. (58) to yield

$$\overline{\ell}_{nn}(k) = 1 + \overline{\ell}. \tag{97}$$

The average number of common neighbors between bottom nodes with hidden variables λ_1 and λ_2 is given by Eq. (51), which now becomes

$$\overline{m}(\lambda_1, \lambda_2) = \frac{Np_0^2}{2\pi}\int_{-\pi}^{\pi} e^{-a\|\lambda_1-\kappa\|^2}e^{-a\|\lambda_2-\kappa\|^2}\mathrm{d}\kappa. \tag{98}$$

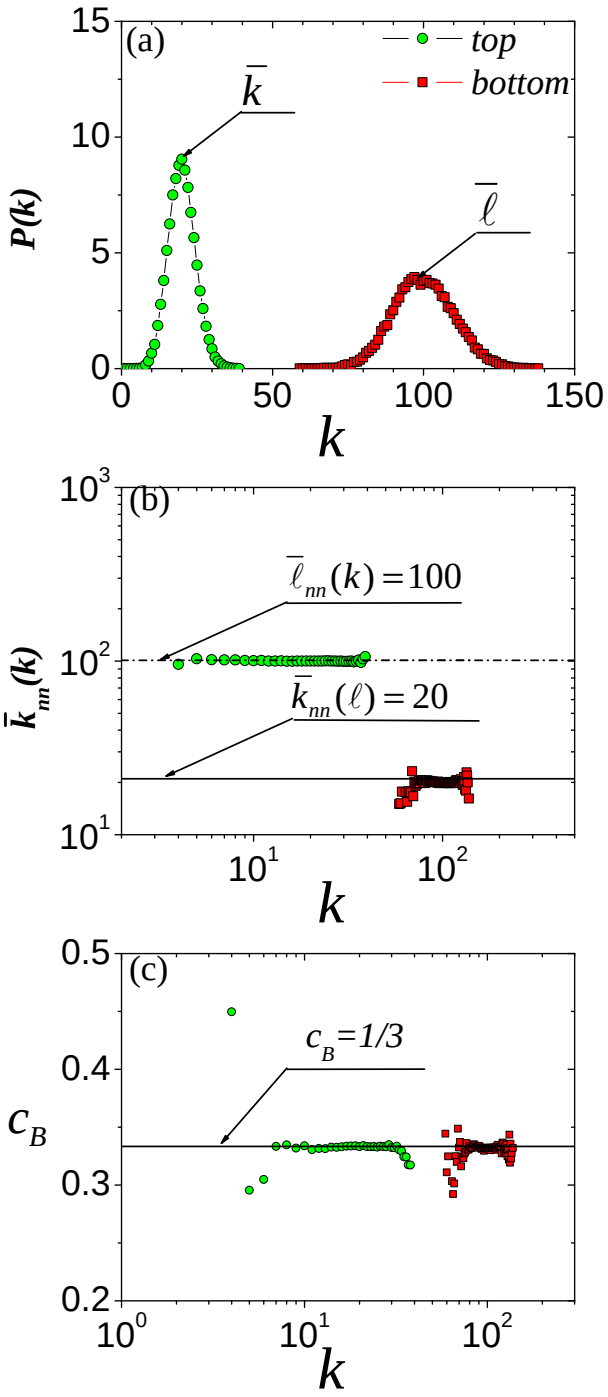

FIG. 4. (Color Online) Stratified bipartite networks. (a) Degree distributions of top (green circles) and bottom (red rectangles) nodes. (b) Average nearest neighbor degrees for top and bottom nodes as a function of node degree. (c) Average bipartite clustering coefficients of top and bottom nodes as a function of node degree. All the plots are for stratified bipartite networks with $N = 10^5$, $M = 2 \times 10^5$, and $\overline{k} = 20$.

To compute $\overline{m}(\lambda_1, \lambda_2)$ we first change the integration variable to $x = \sqrt{a}\kappa$, so that the new integration limits are $\pm\sqrt{a}\pi$. Since $\sqrt{a} \sim M$, in the thermodynamic limit the integration interval becomes $(-\infty, \infty)$, leading

to

$$\overline{m}(\lambda_1,\lambda_2)=\frac{Np_0^2}{\sqrt{8\pi a}}e^{-a\|\lambda_1-\lambda_2\|^2/2}. \quad (99)$$

Inserting the expression for $\overline{m}(\lambda_1,\lambda_2)$ and $\overline{\ell}_{nn}(\kappa)$ into Eqs. (61) and (62) yields the average bipartite clustering coefficient:

$$\overline{c_B}(k)=\overline{c_B}(\kappa)=\frac{1}{3}. \quad (100)$$

To validate the obtained analytical expressions we perform numerical simulations, generating networks with $N=10^5$ and $M=2\times10^5$. To generate a network with a target value of $\overline{k}$ we set $\widetilde{a}$ according to Eq. (91). Figure 4(a) shows the degree distributions for the top and bottom nodes in the model. The degree distributions are well approximated by the Poisson distributions with the averages at $\overline{k}=20$ and $\overline{\ell}=\overline{k}N/M$. Figure 4(b) confirms that there are no correlations: $\overline{\ell}_{nn}(k)$ and $\overline{k}_{nn}(\ell)$ do not depend on node degree, and match Eq. (96). Figure 4(c) shows that clustering is strong, does not depend on either node degree or sizes N, M, and matches the prediction in Eq. (100). The appearance of high bipartite clustering in the stratified model is due to preferential linking of nodes with similar hidden variables.

V. SUMMARY

We have constructed and analyzed a general class of bipartite networks with hidden variables. In this class of bipartite networks, nodes of both type reside in hidden variable spaces, and the connection probability between a pair of nodes is a function of their hidden variables. The independent character of link appearance in the model allows one to calculate analytical expressions for many important topological properties of modeled networks.

The formalism developed here builds up on the hidden variable formalism for unipartite networks [9]. Some basic structural properties of bipartite networks, such as the degree distributions and correlations, are straightforward generalizations of those in unipartite networks. Some other characteristics, such as unipartite projections and bipartite clustering, are unique to bipartite networks.

The hidden variable formalism has proven to be a powerful tool in studying the structure and function of complex networks [22–25]. One particular application of interest for us are network geometry and navigability [26–29]. The formalism developed here can also be useful in inferring individual characteristics, attributes, and annotations of nodes in real bipartite networks.

ACKNOWLEDGMENTS

We thank F. Papadopoulos, M. Á. Serrano, M. Boguñá and kc claffy for many useful discussions and suggestions. This work was supported by NSF Grants No. CNS-0964236, CNS-1039646, CNS-0722070; DHS Grant No. N66001-08-C-2029; and by Cisco Systems.

[1] G. Uchyigit, and M. Y. Ma. *Personalization Techniques and Recommender Systems*, (World Scientific, Singapore, 2008).
[2] J. J. Ramasco, S. N. Dorogovtsev, and R. Pastor-Satorras, Phys. Rev. E. **70** 036106 (2004).
[3] H. Ma, A.-P. Zeng, Bioinformatics **19** (2): 270 (2003).
[4] E. Davidson, and M. Levin, PNAS **102** (14) 4935 (2005).
[5] A. Iamnitchi, M. Ripeanu, I. Foster, INFOCOM (2004).
[6] E. Burgos, H. Ceva, L. Hernandez, R. P. J. Perazzo, M. Devoto, D. Medan, Phys. Rev. E **78**, 046113 (2008).
[7] M. Latapy, C. Magnien, and N. Del Vecchio, Social Networks, **30** (1) 31 (2008).
[8] J.-L. Guillaume, M. Latapy, Inform. Process. Lett. **90** 215 (2004).
[9] M. Boguñá, R. Pastor-Satorras, Phys. Rev. E **68**, 036112 (2003)
[10] D. S. Callaway, J. E. Hopcroft, J. M. Kleinberg, M. E. J. Newman, and S. H. Strogatz, Phys. Rev. E **64**, 041902 (2001).
[11] R. Pastor-Satorras, A. Vazquez, and A. Vespignani, Phys. Rev. Lett. **87** 258701 (2001).
[12] M. E. J. Newman, Phys. Rev. Lett. **89** 208701 (2002).
[13] D. J. Watts, and S. Strogatz, Nature **393** 440 (1998).
[14] P. Zhang, J. Wanga, X. Lia, M. Lia, Z. Dia, and Y. Fana, Physica A **387** 27 6869 (2008).
[15] P. G. Lind, M. C. Gonzalez and H. J. Herrmann, Phys. Rev. E **72** 056127 (2005).
[16] P. Jaccard, Bulletin de la Société Vaudoise des Sciences Naturelles **37** 547 (1901).
[17] H. S. Wilf,*Generatingfunctionology*, 2nd ed. (Academic Press, San Diego, 1994).
[18] E. A. Leicht, P. Holme, and M. E. J. Newman, Phys. Rev. E **73** 026120 (2006).
[19] L. A. Adamic, and E. Adar, Social Networks, **25** 3 (2003).
[20] B. V. Gnedenko, *The theory of probability* (Chelsea, New York, 1962).
[21] Z. Burda, and A. Krzywicki, Phys. Rev. E **67** 046118 (2003).
[22] D. Garlaschelli, A. Capocci, and G. Caldarelli, Nature Phys. **3** 813 (2007).
[23] A. Fekete, G. Vattay, and M. Pósfai, Phys. Rev. E **79**, 065101(R) (2009).
[24] D. Garlaschelli, and M. I. Loffredo, Phys. Rev. E **78**, 015101(R) (2008).
[25] G. A. Miller, Y. Y. Shi, H. Qian, and K. Bomsztyk, Phys. Rev. E **75**, 051910 (2007).
[26] M. Boguñá, D. Krioukov, and kc claffy, Nature Phys. **5**, 74 (2009).

[27] M. Boguñá, and D. Krioukov Phys. Rev. Lett., **102** 058701, (2009).
[28] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguna, Phys. Rev. E. **82**, 036106 (2010).
[29] M. Boguna, F. Papadopoulos, and D. Krioukov, Nature Comm. **1**, 62 (2010).

Appendix A: Derivation of the bipartite clustering coefficient

Here we provide the detailed derivation of the bipartite clustering coefficient defined in Eq. (60). We estimate the average bipartite clustering coefficient of a node with hidden variable κ by calculating the ensemble averages of the numerator and the denominator in Eq. (60):

$$\overline{c_B}(\kappa) = \frac{\langle \sum_{j>l}(m_{jl}-1)\rangle}{\langle \sum_{j>l}(k_j+k_l-m_{jl}-1)\rangle}. \tag{A1}$$

We first focus on the numerator in Eq. (A1):

$$\langle \sum_{j>l}(m_{jl}-1)\rangle = \frac{1}{2}\sum_k g(k|\kappa)k(k-1)\sum_{\lambda_1,\lambda_2}\rho(\lambda_1|\kappa)\rho(\lambda_2|\kappa)\sum_m (m-1)P_{\lambda_1,\lambda_2}(m-1), \tag{A2}$$

where $g(k|\kappa)$ is the κ-to-k propagator, $\rho(\lambda_1|\kappa)$ is the conditional probability that a bottom node has hidden variable λ_1 provided it is connected to a top node with κ, and $P_{\lambda_1,\lambda_2}(m-1)$ is the probability that two bottom nodes with λ_1 and λ_2 have exactly $m-1$ common neighbors besides i. Equation (A2) simplifies to

$$\langle \sum_{j>l}(m_{jl}-1)\rangle = \frac{1}{2}\langle k(k-1)\rangle_\kappa \sum_{\lambda_1,\lambda_2} P(\lambda_1|\kappa)P(\lambda_2|\kappa)\overline{m}(\lambda_1,\lambda_2). \tag{A3}$$

Next we compute the denominator of Eq. (A1):

$$\langle \sum_{j>l}(k_j+k_l-m_{jl}-1)\rangle = \langle \sum_{j>l}(k_j-1+k_l-1)\rangle - \langle \sum_{j>l}(m_{jl}-1)\rangle. \tag{A4}$$

Sum $\langle \sum_{j>l}(m_{jl}-1)\rangle$ is the same as in the numerator, so that we only need to compute $\langle \sum_{j>l}(k_j-1+k_l-1)\rangle$:

$$\sum_{j>l}(k_j-1+k_l-1) = (k_i-1)\sum_{j=1}^{k_i}(k_j-1) = (k_i-1)k_i(k_j-1), \tag{A5}$$

where k_i is degree of node i. Therefore,

$$\langle \sum_{j>l}(k_j-1+k_l-1)\rangle = \sum_k g(k|\kappa)(k-1)k\sum_\lambda \rho(\lambda|\kappa)\sum_\ell(\ell-1)f(\ell-1|\lambda) = \langle k(k-1)\rangle_\kappa \overline{\ell}_{nn}(\kappa). \tag{A6}$$

Using Eqs. (A3) and (A6) we finally obtain

$$\overline{c_B}(\kappa) = \frac{\sum_{\lambda_1,\lambda_2}\rho(\lambda_1|\kappa)\rho(\lambda_2|\kappa)\overline{m}(\lambda_1,\lambda_2)}{2\overline{\ell}_{nn}(\kappa) - \sum_{\lambda_1,\lambda_2}\rho(\lambda_1|\kappa)\rho(\lambda_2|\kappa)\overline{m}(\lambda_1,\lambda_2)}. \tag{A7}$$